PEER REVIEW HISTORY

This paper was submitted to the JECH but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Risk of bias in industry-funded oseltamivir trials: comparison of core reports versus full clinical study reports
AUTHORS	Heneghan, Carl; Jefferson, Tom; Jones, Mark; Doshi, Peter; Del Mar, Chris; Hama, Rokuro; Thompson, Matthew; Onakpoya, Igho

VERSION 1 - REVIEW

REVIEWER	Gerta Rücker, Research Assistant Institute for Medical Biometry and Statistics Medical Center - University of Freiburg Germany
REVIEW RETURNED	04-Apr-2014

GENERAL COMMENTS	Description of the manuscript This manuscript is another one in the sequence of papers the authors published on their experiences when conducting their Cochrane review and updates on Neuraminidase inhibitors ("Tamiflu") for preventing and treating influenza. As known (references 2, 8, 9), these experiences with a manufacturer (Roche) withholding study reports led them to a very critical attitude against the credibility of systematic reviews based on incomplete evidence. Similar experiences were published by others (e.g., references 19, 20 or the articles by Erick Turner [1]). In the present manuscript, the authors originally aimed at comparing three sources of evidence with respect to their information content concerning potential bias: published papers, core reports and full clinical study reports (ordered by increasing detail of information). One problem that occurred was that they could not carry out a comparison of risk of bias judgments of journal publications with core reports or full clinical study reports because they had only secondary publications of the trials. Core reports for 14 trials contained in 10 clinical study reports from 2011 were available, out of 77 full clinical study reports shared later on (April 2013) by the manufacturer. The additional trials were not

	used for this paper (as they wanted to compare the risk of bias assessment with and without full knowledge, if I understand this correctly). The autors decided to skip the level `unclear' from the bias tool when assessing the full reports. The main result was that in all cases where risk of bias was classified `high' based on the core report, it remained `high' for the full report, and for most cases risk classified `low' or `unclear' based on the core report became `high' for the full report. In short, increasing information did not decrease risk of bias. The authors used and extended the Cochrane risk of bias tool. This led them to become quite harsh on the Cochrane risk of bias tool. They write that important elements of bias are not captured by the tool and conclude that (Abstract) `the current Cochrane risk of bias tool is primarily designed to aid the critical evaluation of trials published in journal publications, but full clinical study reports allow for bias to be actually measured rather than reported as an un-quantified risk' and that `further development may be necessary'. The additional items of the extended tool referred to the importance of the trial timeline; they observed, e.g., - whether the trial protocol predated the beginning of participant enrolment, - whether the statistical analysis plan predated participant enrolment, or - when the trial was unblinded. Further, they argue (page 9) that 'Cochrane reviews should increasingly rely on clinical study reports as the basic unit of analysis' and that focus should shift from `risk of bias' to bias itself. Recommendation The authors' experiences in this great deal of work are presented in much detail. As the Cochrane risk of bias tool is widely used and continuously evolving, critique is welcome and the methods and experiences should be made accessible to systematic reviewers both inside and outside Cochrane. Details line 41: Grammar (`Comparison with journal publications was not possible because of publication bias the limits of the Cochrane tool.') - replace `the' with `that' and delete `of'? Or delete `the' and delete `of'? Or is `limits' a substantive?  line 91: Explain what the `A' in the acronym IMRAD means (`Appendix'?)

	lines 126-128 and lines 197-198: Why different order of trials? line 180: `this paper' = the present manuscript?  lines 246-247: `instances where our expectations of having all relevant and consistent information available for our reviews' (grammar:) verb missing Reference [1] Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R.Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med. 2008 Jan 17;358(3):252-60. doi: 10.1056/NEJMsa065779.

REVIEWER	Eveline Nüesch CTU Bern, Department of Clinical Research, University of Bern, Switzerland
REVIEW RETURNED	08-Apr-2014

GENERAL COMMENTS	The authors address the question how the level of detail available for trials influences the assessment of the items in the Cochrane risk of bias tool using a rather unique source of journal reports, core reports and clinical study reports (CSR) from oseltamivir trials. I think an important message of their study is that an increasing level of detail in available reports from trials can affect the risk of bias assessments, and might increase complexity due to large amounts of information, which is potentially not consistent. I have a couple of more detailed comments as outlined below. 1. Can you please clarify your statement in the abstract that "comparison with journal publications was not possible because of publication bias the limits of the Cochrane tool"? (The sentence is not complete.) Where there no journal publications of oseltamivir trials? If so, can you please make it clear that the manuscript is based on comparison between core reports and CSR, but not between published and unpublished reports? 2. Even though CSR provide the most extensive information about a clinical trial, assessment of risk of bias based on CSR is still based on what is reported therein and not a measure of the actual bias in the study. Please clarify. 3. Can you please clarify why there is a higher number of unclear risk of bias assessments based on full CSR (49%) than based on core reports (32%)? It is counterintuitive why a higher level of detail in reports should result in more unclear assessments. 4. You pointed out that there is potential conflicting information in different reports of the same study. Thus, I wonder how you dealt with this in your risk of bias assessments.

	5. It would be very helpful for readers if you could describe the six items of the standard risk of bias tool and provide your criteria for assessment of the individual items (yes, no or unclear). 6. The risk of bias items were not independently assessed by a second reviewer (only checked), which I consider as a clear limitation, because the assessments are crucial in this manuscripts, and often a matter of debate. 7. It doesn't make sense to me to argue that in CSR there should be no ambiguity in risk of bias assessments, and hence not allowing any "unclear" assessments and forcing the assessors to make a judgment of high or low risk of bias. Even though CSR are very detailed, the possibility of an unclear risk of bias should still be there (e.g. if there is inconsistency regarding this item in different parts of the report). 8. A flow-chart would help readers to understand how you ended up with 14 trials for the risk of bias assessments. 9. Can you provide reasons why some risk of bias assessments were low based on core reports and high based on CSR? This is not intuitive to me and might reflect an inadequate judgment of risk of bias based on less information. 10. It is not clear how you could use the randomisation lists themselves rather than the description of the generation of these lists to assess risk of bias due to inadequate randomisation. 11. The items in the Cochrane risk of bias tool are selected according to empirical evidence of bias in clinical trials. It is not clear why you consider e.g. date of trial protocol, date of unblinding, date of enrolment important to assess risk of bias. 12. Even if you know the amount of dropouts or withdrawals, it is not possible to measure the actual bias that arose from this attrition (because there is not data for these patients). Thus, please clarify this (page 8, lines 270ff). 13. In the Box it is unclear why you judged WV15708 and WV15707 as high risk of bias, even though you state that sufficient details were not provided about randomisation and concealment. These should be assessed as "unclear" risk of bias. 14. Tables 1-3 lump all risk of bias items together. However, it would be interesting to see assessments of the individual items based on core reports and CSR and whether there were any remarkable observations for specific items. E.g. it might be that for items such as blinding you needed more detailed reports to reduce the number of unclear assessments than for others, whereas selective reporting might be an issue if you have only journal publications and not anymore if you have CSR.

REVIEWER	Michele Hamm University of Alberta Canada
REVIEW RETURNED	29-Apr-2014

GENERAL COMMENTS	General Comments

	Thank you for the opportunity to review this paper. This was a well written study comparing risk of bias assessments between multiple sources of information for trials evaluating the use of oseltamivir. The question is an important one, given that many decisions are made based on data in published trials, while clinical study reports may represent a more complete source of information but are not readily available. My specific comments below are related to clarification or elaboration on certain points. Specific Comments Abstract: -Lines 34-35: While the authors have included a post hoc sensitivity analysis with unclear assessments, it seems a bit misleading to say nothing was reclassified as unclear when that wasn't an option in the primary analysis. This also comes up in the Results (lines 207-209). -The last sentence of the "Methods and Findings" paragraph could be reworded to be clearer. -In the last sentence of the abstract, do you mean further development of the RoB tool? Strengths and Limitations: -Throughout the text, starting on the Strengths and Limitations page, it seems like adapting the RoB tool might be one of the study objectives. If this is the case, please include an explicit statement. Results: -Lines 199-206: Could this section be reworded/reordered to be made clearer? How many published trials were available compared to the unpublished study reports? -Lines 210-213: Were the "high" RoB assessments driven by any specific domains? How did your findings compare across the domains of the RoB tool? Discussion: -Of the examples provided to illustrate domains that were assessed as high risk of bias, many were focused on missing or inconsistent information. While I understand the authors' rationale for suggesting that there shouldn't be missing details in the full clinical reports, I would suggest tempering statements indicating that the use of these reports allowed the assessment of "bias" rather than "risk of bias," since these RoB judgments are still based on assumptions and are not founded on complete information. -The authors present an interesting argument regarding focusing on bias and clinical study reports, not risk of bias and trial publications, in future Cochrane reviews. Could you please expand on the feasibility of adapting this kind of model, for example, in terms of the resources required and timeliness of these types of reviews?

REVIEWER	Larissa Shamseer Ottawa Hospital Research Institute, Canada
REVIEW RETURNED	02-May-2014

GENERAL COMMENTS	- Author state that they are studying ROB in trial documents with "progressively greater amounts of information", however, it is unclear at times which documents were assessed. The difference between core reports and clinical study reports is particularly unclear. Can the author please provide a box of definitions clearly defining each of the following terms and stating what information is typically contained within each: a journal publication, a core report, a clinical study report. Later on in the methods section, the latter two terms get confused/lumped together (P6, line 183). This is confusing. - Authors could provide a table listing all identified trials, and which trial document(s) they had available/existed (i.e. primary publication, secondary publication of unpublished primary studies, core report, clinical study report). They could then go through the manuscript and eliminate a lot of redundancies in the text. - Authors identify some major reporting issues, especially pertaining to the level and detail of information provided in clinical study reports vs core reports vs publications. Interestingly, some of the extraction items in Appendix 2 resemble items of the CONSORT 2010 checklist. Authors should consider adding discussion of the potential usefulness of CONSORT in improving information contained in journal publications AND all other trial documents such as core and clinical study reports. Adhering to such minimum reporting standards will facilitate easier judgement of design/conduct/methodological validity. Furthermore, a common criticism/excuse for incomplete information in journal publications is that they do not allow for complete details of trials. This is obviously changing given the advent of online publication, and in fact, following the simple recommendations made in the CONSORT checklist, it is indeed possible for all information needed to judge ROB to be included in a publication. This is separate from the issue of data availability (for which full reports are desirable), but information about essential methods and findings can be easily reported within a journal publication. - I think authors could comment on both types of reporting bias – publication bias and selective/incomplete reporting of information within publications vs. core or clinical reports. - Authors should be explicit about which trial documents (i.e. publication, core report, clinical study report) were included/assessed in each version of the Cochrane review. This is very unclear as is.

	- Authors should consider summarizing what the ROB assessments were in each of the 3 published versions of the Cochrane review. Should state whether ROB changed, given the increasing amount of information you had. Were sensitivity analyses carried out in each review and did anything change from version to version? Specifically, did the conclusions change? These points should be stated in the report. - Authors should state that there are unnecessary deficiencies in reporting in documents that are less than the full clinical study report. If all trial documents described a minimum set of information, as recommended in the CONSORT guideline, this would facilitate accurate risk of bias assessments - The conclusions are not clear – are the authors recommending that the ROB tool be modified or that clinical study reports be sought/included in (Cochrane?) systematic reviews or both? - Would be helpful to know whether the core reports and clinical study reports were sent by Roche or EMA as hard copies or electronically ABSTRACT - change once revisions made to fit other recommendations P2, line 28: Open bracket in this sentence, should likely be 'closed after " randomization lists" P2, line 40: This sentence does not make sense, "Comparison with journal publications was not possible because of publication bias the limits of the Cochrane tool." Likely you mean that comparison of both core reports and full clinical study reports to journal publications was not possible, because not all trials were published. In fact a minority of trials were published, which could probably be emphasized. STRENGTHS and LIMITATIONS P3, line 57: uncertainty of biases is likely not representative, perhaps "uncertainty of bias judgements" or assessments? P3, line 59: Unclear whether this second bullet is a pro or con P3, line 63: why would your inexperience limit findings? Because you may not have found all ROB information even if it was in the report? Pls clarify in the manuscript. P3, line 64: This is neither a strength nor limitation. It is a recommendation. Might remove from here and add to the abstract conclusion. Might also be more specific with the language, i.e. the

	Cochrane ROB tool is not adequate for assessing bias in clinical study reports. It could be improved with the addition of other relevant domains associated (or potentially associated) with bias identified in our study. P3, line 67: clarify whether the tool should be used with core reports as well as clinical study reports. And why not non-industry trials? Because there are no clinical study reports? This should be stated or remove the statement about non-industry as it is rather obvious that a tool cannot be used on a document that does not exist. BACKGROUND P3, line 75: change 'mostly' to 'typically' P3, line 72: Does not read well. "standard items considered critical to trial study design", perhaps something like "essential items pertaining to validity of trial design". P3, line 78: are core reports not also more detailed than journal publications? P4, line 90: In line with other comments, it is unclear where your definition for 'core report' ends and 'clinical study report' begins. Also – how does a core report differ from a journal article? Seems to have the same structure as publications (i.e. IMRAD), so is it just more detailed? P4, line 99: This should not be a standalone paragraph. P4, line 100: cite previous version of Cochrane review – important for readers to know which version and year of each review. There are also 3 versions of this review. They most recent version should now be cited, as appropriate, throughout. P4, line 100: this is extremely unclear and makes NO sense! "Unlike most Cochrane reviews, this review was based only on clinical study reports but because of the lateness of delivery of clinical study reports and our funding timelines, the review update was based only on core reports." First, which review is "this review" (needs citation, presumably the 2012 version) and what documents were included in "this review" – core reports or clinical study reports? The review was not based only on clinical study reports if they did not arrive in time to be included. P4, line 104: remove the word "the" P4, line 109: this is not clear" by comparing reports of the same trial with widely varying level of detail", Perhaps could be consistent with objective regarding progressively more info, i.e., "by

	comparing documents containing increasingly detailed reports for each trial included in our review(s)" METHODS The methods of the review are somewhat incoherent and need a major clean up. If I were not familiar with the Cochrane reviews and methods/difficulties obtaining data, I would not understand the timelines as currently written. While this is an acutely popular/landmark topic, it is unfair to assume that all readers now and in the future will have read the referenced reviews and be knowledgeable about the surrounding issues. For instance, the timeline of the 3 reviews, when additional documents were obtained, and which documents were included/assessed in each review version, could be simply summarized in the first paragraph, or even in a table, rather than scattered throughout the methods section as it is. P5, line 126: a Cochrane review is cited, but it is not the most recent review. It would be good to indicate which version of the review you are referring to. P5, line 128: it is unclear whether it was the 14 core reports or the 10 clinical study reports that were obtained by Roche and EMA for the 2012 review. P5, line 138: Is it unusual? Based on what evidence. If Roche is reporting multiple trials in one report, might other companies be doing the same? Also, the following sentence should be linked by semi-colon rather than a hard stop. P5, line 140: remove the word "Trial". Redundant. We already know that you assessed trials… P5, line 141: Is it typical in Cochrane methodology to use an "external review"? why did an external review assess ROB not review authors? P5 line 142: ROB was not re-extracted, it was re-assessed. Please change this. As well, please indicate WHY ROB was re-assessed rather than using original ratings? Did the new assessments result in different ROB judgments than what was published in the review. P5, line 144: this jumps back in time and seems to be ill-placed. You state that the "time-lock" for the 2012 Cochrane review was 12 Apr 2011, but here you state that you only started receiving appendices P6, line 154: do you mean 'review' and not 'reviews'? singular…? P6, line 169: So was the longer ROB form used to assess both core reports and clinical study reports. There seems to be no mention of

	how core reports were assessed. Was the conventional ROB tool applied? RESULTS P7, line 196: "We had these.."… what are 'these' referring to – core reports or clinical study reports. P7, line 201: your ROB assessments were based on secondary and not primary journal publications – is this because primary journal articles for those trials don't exist? Please make this clear. P15: why haven't authors presented ROB broken down by domain, since ROB is never summed up into a single overall score. It would be more interesting and provide more insight as to what items were better reported/obtained in the larger documents. P17: M2 likely refers to Module 2, but this is not stated anywhere. Ditto for 'SAP' (p19) DISCUSSION - Authors refer to an instrument they developed, which is essentially the additional extraction items which they suggest should be ROB items for clinical study reports. Sch an instrument would need much more evaluation, application, and refinement before suggesting it is as a definitive tool for use in future research. Authors should be more cautious with their recommendations and list these issues. - If authors are going to suggest their instrument to others, they should provide an operational definition of each extraction element, as Cochrane has for each ROB element. - I am not convinced that the instrument should be recommended for future use until more careful consideration and repeated application is carried out in other studies. While it is a great resource, it should likely not be highlighted/recommended as a strength of the study - See also the general comments above This work is obviously important, and no doubt well done, but it is very poorly written, described, and at times seems to be haphazardly thrown together. It would be difficult to publish a study about reporting issues when the study itself is not well reported.

	Comments	Notes	Action
1.	Editor points: 1. You mention that this is the first in a series of papers. Can you please provide an outline for the other papers planned for this series;	We think this may be a misunderstanding. There is no planned series. We are in the process of publishing the updated A159 review in the Cochrane Database of Systematic Reviews and there may be other journal article spin offs, but we have no definite dates at present	Nil
2.	2. Abstract, Methods & Findings. Please include a sentence at the end of this section to describe the main limitations of your analysis (Please note the abstract can be longer than 300 words in the main text file);		The following phrase has been inserted in M&F: One limitation of our study was our relative inexperience in dealing with large information sets. We also found risk of bias judgments to at times be subjective. Also, our focus on industry trials reported in clinical study reports may not apply directly to non-industry trials. We also had no control over which studies EMA provided and cannot exclude selection bias. We could not validly compare risk of bias based on journal publications because our assessments were largely based on secondary (i.e. not primary) publications of the trials and an outdated risk of bias tool.
3.	3. How much disagreement there was between the 3 reviewers and how many items were automatically be categorized as “high” risk of bias due to missing information;	For the second assessment (i.e. of full clinical study reports) we did not have arbitration or a third party judging, as we reached consensus through discussion of a face to face meeting of all investigators.	The text has been amended to read: “These were carried out by a single reviewer, checked by a second with final consensus reached through a face-to-face discussion

			among the entire group".
4.	4. Were the methods for risk of bias judgments from the previous Cochrane reviews [10,11] the same? ie "carried out by a single reviewer, checked by a second, independently judged by a third person";	Yes	The sentence "The extraction and adjudication methods used were the same as those used in our subsequent unified Cochrane review [6]." Has been added to the text.
5.	5. Please include an ethics statement in the methods section;	Could not find a standard statement on the PLoS website, hope the one in next box is acceptable. Please note that Ethical considerations in conducting a Cochrane review are available at: http://www.cochrane.org/editorial-and-publishing-policy-resource/ethical-considerations	Inserted at the end of Methods: "Ethics approval and patient consent forms are not provided as they are not necessary for a Cochrane review, of which this study is a product".
6.	6. Thank you for providing a link to the reports. We very much support the reports being made available and we ask that you deposit the information in DRYAD; http://datadryad.org/	Posting to Dryad requires We support Dryad too but cannot post the clinical study reports as soon as we have a doi.	For peer review purposes we set up a Dropbox access which is still extant: https://www.dropbox.com/s/8ofs5qyjmkd7bxk/Jefferson%20et%20al%20ROB%20paper%20manuscript%20-%20supporting%20files.zip Please note that our intention was stated in our original submission letter ("As a supplementary item we are making the source table of all risk of bias assessments available to you and your reviewers and are happy to share the relevant complete CSRs. This however will have to be through a drop box type system, given their size") and the url was passed on to your editorial

			office on the First of November.
7.	Academic Editor comments: The lack of clarity around which parts of the CSRs were analyzed at which stages of the study needs to be fixed. I'm also concerned about eliminating the "Unclear" category when evaluating complete study reports. The authors should provide a better rationale for that choice, and discuss the implications of that choice for the conclusions they can draw. Alternatively, they can include the "Unclear" category in their evaluations of CSRs and present that data in the revised paper;	Two issues are mentioned in this comment. The 3-stages issue: the convergence of editors and reviewers have led us to reanalyze our publications - core report – full report sequence. We realized that comparing core or full reports with publications across reviews is potentially confusing and misleading. Many of the trials included in the latter Cochrane reviews had no correspondent articles because of sizeable publication bias. The Cochrane Risk of Bias tool changed in 2010 making direct comparison of assessments difficult. Finally, publications reported far fewer items that could be assessed, leaving many of the boxes empty. The rationale for eliminating the "unclear" risk of bias category issue: this seems to us difficult to argue with, if one is to simply skim the clinical study reports we have made available. The reports form the scientific basis for approval of a pharmaceutical. Everything that should be of interest to regulators to assess whether the drug is better than placebo (including thousands of pages of individual listings) should be in the reports. Risk implies uncertainty, but reports are supposed to	We have simplified the stages throughout the manuscript by downsizing the first stage (publications) giving the reasons listed in the column to the left. For example in Results the following has been added: We could not carry out a comparison of risk of bias judgments of journal articles with core reports or full clinical study reports because our assessments were largely based on secondary and not primary publications of the trials and an outdated risk of bias tool. There were therefore too few studies for which we had distinct risk of bias judgments of primary journal publications (many studies for which we have clinically study reports were and remain unpublished). In addition, the current Cochrane risk of bias tool was introduced after the production of our review based on published articles, making the comparison, had we had the data to undertake it, more difficult to interpret and possibly unfair.

		provide certainty of full design and reporting of the trial. If an item is not reported bias is present. We realize that switching gears from publications to clinical study reports is difficult, but it is essential now as we move to a different evidence paradigm. As an example, if an item such as randomization is not reported in full, bias is present.	
8.	Reviewer Notes: Reviewer #1: I confine my remarks to statistical aspects of this paper. These were very simple; I think some tests could be usefully added and I have some other comments as well;	See serial 12	Nil
9.	General comment: Why is risk of bias categorized into "low", "unclear" and "high"? Since the tool used to assess the risk has 7 items, couldn't more precise levels of risk be determined;	Apologies, there is probably a misunderstanding. The Cochrane tool has 6 domains, with possibly more than one source of bias each of which is rated low/unclear/high. So the levels of risk are 3. It is impossible to list all possible sources of “other bias” and the tool does not list them.	The following text has been added to Methods: “The current Cochrane risk of bias tool was first introduced in 2010. The tool consists of six domains, each may have more than one source of bias application, depending on the subject matter.[7] Our applications were as follows: selection bias (random sequence generation and allocation concealment), performance bias (blinding of participants and personnel – all outcomes), detection bias (blinding of outcome assessment - all outcomes), attrition bias (influenza symptoms, complications and harms outcome data), reporting bias (selective reporting) and other bias. The identification of sources of other bias are left at the reviewers’ discretion.”

10.	line 109 ff Say more about how the Cochrane method works. The 7 items are in an appendix, they should be here, as should be a description of how they are combined;	See serial 9	See Serial 9
11.	line 140 ff Why does complete information mean that no trial will have an "unclear" amount of risk of bias;	See Serial 7	Nil
12.	line 151-2 I am certainly not a big advocate of p values, but statistical tests do have some purpose. They could be used here to distinguish the amount of risk of bias at different levels of information. A simple chi-square test would probably suffice. (And would doubtless be hugely sig.);	Serial 8 also refers	A simple chi-square test is not appropriate because we are (re) assessing the same trials based on differing amounts of information available In addition the purpose of conducting a hypothesis testing is to provide inferences to the wider population. However we are unsure if this is appropriate given our lack of experience with clinical study reports.
13.	line 165 Since the authors said that no complete report could be unclear, it is not sensible to report on the proportion that were unclear. It is 0 by definition;		Text has been clarified to: “compared to none using complete clinical study reports”
14.	Tables 1 - 3 would probably be clearer if the columns were ordered "high", "unclear", "low";		Tables 1 and 3 were removed. Table 2 (now 1) has been edited accordingly
15.	Figure 1 would probably be clearer as a mosaic plot (or two);	After removal of the publications ROB analysis from the review (see Serial 7) we played with several figure formats. None were satisfactory and the figure was removed	Figure removed

16.	Reviewer #2: Reviewer: Beate Wieseler	-	
17.	General comment: This study provides important and new information on the relevance of clinical study reports (CSRs). As such it is specifically important given the current initiatives to make CSRs publicly available. It furthermore addresses possible implications of availability of CSRs for systematic reviewers, which might start a timely discussion. However, some key issues would need to clarified before publication (e.g. unclear distinction between a "core" and a "full" clinical study report in the methods section as well as inconsistencies in sample sizes in the results section);	The difficulty is double for those not used to clinical study reports: Roche-speak ("Module 1) and ICH-speak (e.g. "core report").	In the introduction the following has been added …."For the purposes of this paper the core report plus all it appendices (roughly equivalent to modules II to V in oseltamivir clinical study reports) will be known as the full clinical study report"…. We also have gone through the text to ensure consistency of terms
18.	Abstract Lines 34-35: 11 articles, 15 CSRs: how many studies;		The text now reads: "We used and extended the Cochrane risk of bias tool to assess and compare risk of bias of 14 oseltamivir trials (reported in 10 clinical study reports) obtained from the European Medicines Agency (EMA) and its manufacturer, Roche."
19.	Lines 36 - 44: The results are difficult to understand, please also see comments on results section of manuscript		We have deleted references to 2009, 2012, 2013 etc and re-written the results. We hope this is now clear. We now refer to "journal publications", "core reports" and "full clinical study reports" in a consistent manner
20.	Line 43: It is unclear from the abstract how the results presented support the conclusions. From my point of view this only becomes clear		The conclusions have been edited

	from the discussion;		
21.	Introduction Line 69: I would like to suggest using "according to sections 1-15 …" because ICH E3 is not meant as a template of a CSR but describes required content (and the general structure);	ICH E3 is a guideline, and it would be wrong to suggest the wording 'according to ICH E3' in the text, as it is not mandatory for industry to follow	Nil
22.	Line 70-76: The first part of the section introduces a definition of a core CSR (report according to sections 1-15 of ICH E3); however, in the second part of the sentence appendices are included (which would not be part of a core report). This could easily be solved by splitting the sentence in two and clearly referring to appendices as part of a full CSR. The definition of a core and full CSR is also unclear from other parts of the manuscript (please see below);		The following had been added to the Introduction: "For the purposes of this paper the core report plus all it appendices (roughly equivalent to modules II to V in oseltamivir clinical study reports) will be known as the full clinical study report"
23.	Lines 78-85: It remains unclear from the text why funding timelines resulted in only 20 of 32 trials being included and why only core reports were included. This only becomes clearer in the methods section. As such I find this paragraph in the introduction rather difficult to understand. Please rephrase;		The para now reads: "In 2012, we published an update of our Cochrane review of neuraminidase inhibitors for which a total of 32 oseltamivir trials were eligible. Unlike most Cochrane reviews, this review was based only on clinical study reports but because of the lateness of delivery of clinical study reports and our funding timelines, the review update was based only on core reports."
24.	Lines 93-98: What is the research question; 1) Investigating the influence of progressively greater amounts of information in CSRs on the	The 3 objectives are summarised in the abstract phrase: "Here we analyze whether progressively greater amounts of information	The end of the Intro now reads: "In this report we describe our use of these tools

	risk of bias assessment or;	and detail in clinical study reports….”	to address three specific questions: 1. Do core reports change the risk of bias evaluation compared to published papers? 2. Do full clinical study reports change the risk of bias evaluation compared to core reports? 3. Do full clinical study reports change the risk of bias evaluation compared to published papers? In summary we intended to analyze whether progressively greater amounts of information and detail in clinical study reports (including trial protocols, statistical analysis plans, certificates of analyses, individual participant data listings and randomization lists) affected our risk of bias assessments”. We then explain in the results that objective 3 could not be achieved (see also serial 7)
25.	2) Investigating whether the Cochrane risk of bias tool plus the additional instrument can be used to assess risk of bias of trials reported in CSRs;		
26.	Methods General comments on methods: The study analyses the outcome of a risk of bias assessment based on journal publications, core CSRs and complete CSRs. Lines 68 to 70 of the manuscript define the core CSR as those parts of the CSR prepared according to sections 1-15 of ICH E3, i.e. the main report from the title page to the reference list but without the appendices. Using the Roche CSR structure this would be Module I only. From the methods section it remains unclear whether this definition of a core report was	On reflection we agree with comments.	We have taken the unsatisfactory comparison with risk of bias of journal publications and simplified the text comparing core reports with full reports for the 14 trials in 10 CSRs mentioned at the beginning of Methods. See also Serial 7

	also used for the analysis or whether any appendices (either received from EMA or provided by Roche, Module II according to Roche CSR structure) were also included in the assessment of risk of bias labeled as a core report in Tables 1 and 2. Please clarify. Please also provide a clear definition of core and full CSR in the methods section. Although there is no established definition in the literature, from my point of view the core report would be the part of the CSR according to sections 1-15 of ICH E3, i.e. without any appendices;		
27.	Please also consider the following two issues: 1) From the supporting material it seems that appendices (Module 2) were included in the "core report" in your analysis For example, in the Excel table provided as supporting information (CIST M2 table) line 36 for study WP16263 (element: blinding of participants) describes as the rationale for the 2012 assessment (which I understand is the assessment you used as "based on core CSR", lines 107-108 of the manuscript): "Placebo and oseltamivir capsules were described as having non-identical appearances from the certificate of analysis: oseltamivir: "Body: grey, opaque; cap: light yellow, opaque" placebo: "Body: grey, opaque; cap: ivory, opaque""; According to the definition of a core report referring to sections 1-15 of ICH E3, the certificates of analysis would not be part of the	We agree with the comments and have taken the action listed in serials 7 and 26. Placebo description was available in the text of one core report	

	core report because they are part of the appendices;		
28.	2) According to the Jefferson 2012 Cochrane Review (Table 9), more than Module 1 seems to have been available for at least 5 CSRs. Does that mean that more than Module 1 was used in the risk of bias assessment in the Jefferson 2012 review (and is the risk of bias assessment of Jefferson 2012 indeed presented as an assessment "based on core CSRs" in the manuscript?);	Jefferson 2012 used mostly Ms1, but also some Ms2 as they came in before timelock. For this paper, we're restricting our analysis to trials for which we only had M1 in Jefferson 2012. The order in which we received clinical study reports was outside our control. It could have introduced some bias, although we have no evidence of that.	This reflection has been added as a potential limitation of our study
29.	Lines 100-103: It is unusual that several studies are reported together in one CSR. Please explain in more detail (what were the reasons for this, was this justified, did the reports still include a full account of information on the individual studies?)		The following has been added to Methods: "The reporting of more than one trial in the same clinical study report is unusual. Roche gave low influenza circulation and the consequent need to pool studies as the reason."
30.	Line 104: Since there are a number of Cochrane reviews on neuraminidase inhibitors, please include the relevant citation: … timelock for our 2012 Cochrane review update [reference];		References have been added
31.	Lines 109-120: It does not become clear from this paragraph for how many studies core reports or core reports with (exactly which) appendices were available. Please clarify. An appendix table describing the available parts for each of the 15 reports might be helpful;	NB: **Reference 10** in the manuscript is a bookmark for the reference of our latest Cochrane review update which is currently undergoing peer review. We will insert the full reference as soon as it format is known.	We have revised the text to as follows: "In April 2011, we began to obtain the appendices of the clinical study reports included in our review. For most clinical study reports we

			requested, EMA had the protocol, protocol amendments, statistical analysis plan, blank case report forms, and other appendices contained in what Roche terms the second "module" of a full clinical study report (see Appendix 1). However EMA did not possess—and therefore could not provide us with—full clinical study reports with the exception of trial WP16263.[8] For approximately three years Roche had repeatedly refused our requests for full clinical study reports.[9] In the course of carrying out these new extractions, Roche changed its policy on access to data and pledged in April 2013 to share with us 74full clinical study reports (www.bmj.com/tamiflu/roche). Twenty trials were included in the analysis of our current Cochrane review.[10]. As we were already in possession of core reports and appendices such as the protocol and statistical analysis plan for the 14 trials in this analysis, the additional data for other clinical study reports provided by Roche does not concern this paper. In the Clinical study reports Roche redacted information that they judged to be of "legitimate commercial interest" or present a risk of trial participant re-identification. For our purposes, the redactions did not impede an analysis of risk of bias. Based on our growing familiarity with clinical study reports, we designed and piloted an extraction sheet to record how our understanding of the trials changed in light of

			availability of the additional appendices. “
32.	Lines 128-133: In contrast to lines 100 to 103 here you are referring to 74 CSRs. Please clarify the different numbers. Which CSRs (studies) were used in your study (only the 15 reports [covering 20 studies] mentioned in lines 100-103?);		Text has been clarified throughout the manuscript
33.	Lines 135-139: Which studies were included in your investigation presented in the paper;		See Methods and response to Serial 31
34.	Lines 140-146: For assessments of risk of bias based on full CSRs you did not allow an "unclear" judgment. This seems to be justified. However, since e.g. unclear allocation concealment is not necessarily inappropriate allocation concealment, you might be underestimating the quality of the trials. This has implications for the interpretation of your results, which does not become fully clear from your results text and discussion. Please address this issue (please also see comments below);	We think the reviewer is taking the risk of bias judgments as 100% objective. All such judgments are challengeable. If allocation was unclear, we would have said “high” risk of bias, because of the logic of our judgment already explained. When the originals trials were designed, allocation concealment would not have been seen as a source bias. Initial publications that revealed its importance only came to light in 1995 and its incorporation into CONSORT was much later than the trial ending (JAMA. 1995 Nov 8;274(18):1456-8.Subverting randomization in controlled trials. Schulz KF et al)	Nil
35.	Lines 149-150: It does not become clear how the extraction of risk of bias assessments described in lines 149 to 150 (citing the 2010 [Jefferson et al] and the 2012 [Wang et al] Cochrane review) relates to lines 107-108 (citing the 2012 Jefferson et al review). Please clarify. It might be meaningful to describe all	Extraction is too technical a term	Edited to: “we used” instead of “we extracted”

	data extractions together in one paragraph;		
36.	Results General comment on results: The timing of risk of bias assessments in the various versions of the Cochrane reviews and the documents on which these assessments were based does not become clear from the manuscript without going back to the cited Cochrane reviews and the supplementary materials. In addition, the number of studies and publications, core reports and full CSRs used is unclear from the text and tables. Specifically, it is not clear why 11 core reports are used in Table 1, 15 in Table 2 and 11 in Table 3. Please clarify. It might be helpful to provide a flowchart or some other sort of graphical representation;		See responses at serials 7 and 26
37.	General comment on tables: 11 core reports are used in Table 1, 15 in Table 2 and 11 in Table 3. This probably is due to the fact that you are trying to include the maximum of available risk of bias assessments in your analysis. As an additional analysis, it might be useful to perform all comparisons on a consistent sample of trials/documents. This would allow describing information gain along the line of adding more and more parts of CSR documents in a given sample of trials (please see also comment on lines 159-161 below);	Good suggestion	We have restricted our comparisons to 14 trials throughout (core vs full study reports) See also responses at serials 7 and 26
38.	Tables 1 to 3: Please explain how the total number of judgments is derived in the tables	We agree	Only 1 Table in the manuscript now

	(89 in Table 1, 130 in Table 2, 90 in Table 3). I assume that, in addition to the different number of core reports included, the differences are due to different numbers of outcomes assessed. However, it remains unclear why you have a different number of judgments in the publications sample in Table 1 (89) and Table 3 (90);		
39.	Table 2: Compared to the judgments from core reports, adding information from complete CSRs did not change any "high" judgments. According to your methodology you do not accept any "unclear" judgment any more at this stage. Therefore, any "unclear" judgments that cannot be solved from complete CSRs change into "high" judgments. It would be interesting to know which part of the "high" judgments is due to additional information from complete CSRs leading to informed "high" judgments and which part is due to still unclear information;	Good point. NB: **Reference 10** in the manuscript is a bookmark for the reference of our latest Cochrane review update which is currently undergoing peer review. We will insert the full reference as soon as it format is known.	We have inserted a sensitivity analysis in the text of results: “Had we kept the unclear risk of bias judgment in our current review [10] we would have had 64 unclear occurences. The breakdown of these 64 into the various attributes is: Attrition bias: symptoms (10); complications (9); safety (15) – these are unclear because we do not know the impact of missing symptoms data; unclear definitions for complications; and compliharms Other bias (13) – these are unclear due to the unknown effect of the de-hydrochloric acid Performance bias (6) – these are unclear due to the missing certificate of analysis describing the placebo appearance Selection bias (10) – these are unclear due to the missing or unclear randomisation lists meaning we cannot confirm random sequence generation

			Detection bias (1) – unclear due to unknown impact of different coloured placebo caps on outcome assessment” We have inserted two new tables reporting the results of a sensitivity analysis allowing unclear judgments in both core and full clinical study reports (Table 2 and 3).
40.	Figure 1: Please clarify in the figure legend why the columns for 2012 and 2013 do not present all low, unclear and high judgments together;		Figure 1 has been deleted
41.	Lines 158: Table 1 includes 74/89 "unclear" judgments for journal publications, not 75/90, please clarify;		74/89 is correct
42.	Lines 159 - 161: I don't think one can follow the "unclear" judgments from Table 1 to Table 2 because there are different samples in the two tables (you might want to add a set of tables based on the same sample to allow for such a follow-up from publication to core report to full report). It would be interesting to provide examples of changes from unclear to low to high describing the information which led to these changes in judgment;	.	See serial 39, but remember the judgments are subjective
43.	Lines 162-167: The fact that there are 0 judgments of unclear risk of bias based on full		See serials 39 and 42 and clarifications inserted

	CSRs is (just) a consequence of your definition. Please clarify this in the text of the results section; otherwise the reader might understand that full CSRs provided full information in all cases. Furthermore, please clarify how many of the "high" judgments are due to still unclear information and how many are due to the assessment of the available information in the full CSRs;		in text
44.	Lines 154-167: It would be interesting to know in which of the elements of the risk of bias assessments the most prominent changes were seen.Is it possible to at least qualitatively describe what type of information led to the changes;		See serials 39 and 42 and clarifications inserted in text
45.	Discussion Lines 184-185: ""Unclear" risk of bias often became a more certain "low" or "high" risk of bias, or even certainty of bias or certainty of absence of bias. I understand that more complete information from a full CSR can result in certainty of bias or certainty of absence of bias. However, from my point of view it is open for discussion whether there is a higher risk of bias from the fact that there is missing information from a full CSR versus a core CSR or a (full or core) CSR versus a journal publication. E.g. with examples 1and 2 from Box 1: missing information from the core CSR leads to unclear risk of bias while missing information from a		See serials 39 and 42 and clarifications inserted in text

	full CSR leads to high risk of bias. In your methods section you define missing information from a CSR as high risk of bias. You might consider discussing this in more detail;		
46.	Lines 186-187: "Certainty or low levels of uncertainty are due to our expectations regarding the complete clinical study reports." I understand you are expecting certainty or low levels of uncertainty from full CSRs. However, I do not understand what this sentence means.;	Thank you	Text changed to: ” When the information was not available, our judgments changed because we found gaps in the availability of information and inconsistent information. Whether the full study reports represent an exhaustive and coherent source of trial narrative and data remains unclear”.
47.	Potential additional points for the discussion: You correctly point out that with full CSRs there is the possibility of understanding the timeline of a study concerning study planning, enrolment and treatment of patients and analysis and reporting of data. According to my experience, there might be changes to the protocol after the start (or even the end) of enrolment (but before unblinding). I am not sure this always results in a high risk of bias (in blinded studies). As with the availability of full CSRs, systematic reviewers would use this information in their risk of bias assessment, there is a need to discuss under which circumstances this should result in a high risk of bias and in which circumstances this would be less critical. The same is true for the development of a full statistical analysis	The reviewer appears to regard risk of bias assessments as objective and assigned on the basis of pre-set scenarios. This is not so especially since this is the first time to our knowledge that such assessments are made on clinical study reports. We find it very difficult to speculate on specific scenarios of levels of bias. This is why we agreed to regard any unclear item in full study reports as unclear.	Nil

	plan (for blinded studies). While obviously the crucial issues and general lines of analysis should be pre-defined in the protocol, often detailed analysis plans are prepared while the study is ongoing (but before unblinding). This need for clarification (and other open questions resulting from access to full CSRs for risk of bias assessment) could be an additional point for the discussion.;		
48.	I am wondering if the discussion could also address earlier work on the relevance of study protocols for risk of bias assessment and what full CSRs would add. For example a paper by Soares et al. addresses information gain from protocols on study methods: Soares H et al. Bad Reporting does not mean bad methods for randomized trials: observational study of randomized controlled trials performed by the Radiation Therapy Oncology Group. BMJ 2004; 328:22 A commentary on the Soares paper is also addressing the impact for systematic reviews: Del Giblio A et al. Commentary: The quality of randomized controlled trials may be better than assumed. BMJ 2004;328:24 We have quantified information gain from CSRs for some methods items used for risk of bias assessments: Wieseler et al. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports and journal publications. BMJ 2011;	We are not sure the Soares paper is an appropriate example as it compared methodological features in protocols and published counterparts in a population of 56/59 published trials. We have already referred to Wieseler's later study (ref 20): Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervölgyi V, Kohlepp P, et al. 395 Completeness of Reporting of Patient-Relevant Clinical Trial Outcomes: Comparison 396 of Unpublished Clinical Study Reports with Publicly Available Data. PLoS Med. 2013 397 Oct 8;10(10):e1001526.;	Nil

	344:d8141;		
49.	Comment on supporting materials Excel table provided as supporting information (CIST M2 table): The study IDs given in the table do not match the study numbers given in the methods section (which probably are report numbers). Please clarify by providing both study IDs and report numbers in the appendix table.;	The Table reports risk of bias assessments for all trials in the reviews, not just the ones in this study	Nil
50.	Reviewer #3: This is an excellent summary of the changes in the risk of bias comparing the data in published articles in journals to clinical study reports. The information is timely and useful, and points to the need for examining all the evidence when evaluating medical interventions.;	Thank you	
51.	A few comments that would aid in clarity of the manuscript: * The number of actual clinical trials vs the number of complete study reports is not clearly explained up front and it take a while to gather that there are multiple studies included in the various study reports. Would be helpful to clearly explain this in the abstract and introduction;	We agree	See serials 7 and 26
52.	* The point about the Cochrane risk of bias tool not being optimal is an important one. The Cochrane tool has a broad category called "other risk of bias". Such a broad category is often not helpful in examining data as it relies	We agree	See serial 9

	on the reviewers personal knowledge of types of bias or actually looking for various types of bias. The seminal paper by Sackett listed over 30 types of bias that could be present in various types of studies (both randomized and observational) and represents a more comprehensive way of evaluating bias;		
53.	* One major issue is the authors eliminating the category of "unclear" risk of bias when evaluating complete study reports. While it is agreed that there should be no information left out of these reports, the unfortunate truth is that even with these more thorough sources of information, sufficient detail is often missing. For instance in Box 1 the authors point out that insufficient detail is available regarding issues like the randomization code. Therefore there is still "unclear" risk of bias when information is missing. Eliminating the "unclear" category has two consequences: first it means that the published and the complete study reports are not judged by the same standard which could itself bias the assessments of bias; second, it assumes that absence of evidence if evidence of absence - because detail is not present then it must not have been done or been done wrong. This is in unverifiable assumption. It would be better to reanalyze the results and use the "unclear" category when insufficient detail is present in CSRs. This is still a major problem in terms of trial transparency since as the authors state there is no reason these	We agree	See responses to serials 7 and 39

	important details should be withheld.;		
54.	* The idea that missing data results in actual bias rather than "risk of" bias is an interesting notion. However in some situations when analyzing different imputation methods the best and worse case scenarios give the same qualitative (if not quantitative) results so that the bias does not affect the robustness of the conclusions. However the authors point is well taken that in many cases real bias exists and it more than just "potential" bias.;	Thank you	Nil
55.	* It would be helpful if the authors could make some suggestions as to when bias is so limiting as to make the conclusions of meta-analyses unreliable. It is unfortunate to see meta-analyses in the literature of studies with high risk of bias, but when the results of these meta-analyses are presented the risk of bias is not considered or ignored, and the conclusions are presented as confirmatory.	We agree with the reviewer. We think when bias is so limiting as to make a meta-analysis unreliable, it either should not be done or should be done alongside a prominent explanation of the extreme limitations.	The following text has been inserted in Discussion: “We suggest that when bias is so limiting as to make meta-analysis results unreliable, it either should not be done or a prominent explanation of its clear limitations should be posted alongside the meta-analysis.”

VERSION 1 – AUTHOR RESPONSE

VERSION 2 – REVIEW

REVIEWER	Larissa Shamseer Ottawa Hospital Research Institute
REVIEW RETURNED	22-Jun-2014

GENERAL COMMENTS	Overall, I think authors have done a satisfactory job of addressing most of the comments from all reviewers and I thank them for doing that in such an organized manner. There are a couple unaddressed/dismissed points that I don't think authors considered carefully enough and I am bringing them to attention again here (Feel free to include anything from the text below in your manuscript). I understand that the point authors are trying to make is that clinical study reports contain more information than core reports, both also contain more information than journal publications, when published. This latter point exemplifies the problems of both incomplete reporting of published research and non-publication of trials. However , I see from their responses that authors do not think it is important to discuss either the problems with the completeness of reporting (as per CONSORT) or publication bias. Even if these weren't the main issues, they still pertain to authors' experiences with Tamiflu. I think it is a missed opportunity not to comment on them here. Doing so will tie your findings back to the broader context of well-known reporting problems; your findings about ROB exacerbate current reporting problems and contribute to existing literature. Along the same lines, authors suggest that obtaining core reports and clinical study reports would the best approach for systematic reviewers to do going forward ("we can think of no alternative to the use of full clinical study reports"??) . However, obtaining full clinical study reports is not exactly a feasible proposition for future review authors given current regulations (or lack thereof), as authors know from their own experience (how many years did Tamiflu take? 6 months doesn't seem like a realistic approximation). There is a lesser, but popular alternative right now: Why not use this microphone (i.e. publication) as a call for both companies carrying out trials and individual trialist to make it a priority to both publish a complete and transparent account of their research, accompanied by trial data. You're making it seem like the onus is solely on reviewers to obtain data, however primary researchers should be providing/publishing all necessary data in the

	first place, in order to facilitate syntheses and replication. TITLE Two of your study questions describe comparison of journal publications to core and clinical study reports, why was the title changed only reflect the latter two? Shouldn't it be a comparison of journal publications, and unpublished core reports and clinical study reports? ABSTRACT In the Methods, this should be modified to make it clearer what was being compared, "With more detailed information documented in full clinical study reports for each trial…" The conclusion is not very clear – which document type had a lower ROB? Be explicit. In the first sentence, perhaps something like "Full clinical study reports were consistently assessed as having a high ROB compared to core reports (and journal publications where available) which were previousy assessed as 'unclear' ". Also the next sentence in the conclusion is not true: "This may mean risk of bias has been insufficiently reported in other Cochrane review assessments limited to published research." The reporting of ROB is not an issue, it's assessment due to reporting issues in the primary research. This is an important finding criticizing Cochrane, I would be more careful about wording. Might instead read something like "Assessments of risk of bias in Cochrane reviews are likely underestimated since they almost exclusively rely on published primary research which is not neither always available or, when it is, inadequately reported. Page 8, line 242: You state there being too few journal publications available to you to complete assessments with, but can you just state how many to make it easy for the reader? Was it that 5 trials were published? It's not clear from what you've written.

VERSION 2 – AUTHOR RESPONSE

	Comments	Notes	Action
56.	Editor points: 1. You mention that this is the first in a series of papers. Can you please provide an outline for the other papers planned for this series;	We think this may be a misunderstanding. There is no planned series. We are in the process of publishing the updated A159 review in the Cochrane Database of Systematic Reviews and there may be other journal article spin offs, but we have no definite dates at present	Nil
57.	2. Abstract, Methods & Findings. Please include a sentence at the end of this section to describe the main limitations of your analysis (Please note the abstract can be longer than 300 words in the main text file);		The following phrase has been inserted in M&F: One limitation of our study was our relative inexperience in dealing with large information sets. We also found risk of bias judgments to at times be subjective. Also, our focus on industry trials reported in clinical study reports may not apply directly to non-industry trials. We also had no control over which studies EMA provided and cannot exclude selection bias. We could not validly compare risk of bias based on journal publications because our assessments were largely based on secondary (i.e. not primary) publications of the trials and an outdated risk of bias tool.
58.	3. How much disagreement there was between the 3 reviewers and how many items were automatically be categorized as “high” risk of bias due to missing information;	For the second assessment (i.e. of full clinical study reports) we did not have arbitration or a third party judging, as we reached consensus through discussion of a face to face meeting of all investigators.	The text has been amended to read: “These were carried out by a single reviewer, checked by a second with final consensus reached through a face-to-face discussion

			among the entire group".
59.	4. Were the methods for risk of bias judgments from the previous Cochrane reviews [10,11] the same? ie "carried out by a single reviewer, checked by a second, independently judged by a third person";	Yes	The sentence "The extraction and adjudication methods used were the same as those used in our subsequent unified Cochrane review [6]." Has been added to the text.
60.	5. Please include an ethics statement in the methods section;	Could not find a standard statement on the PLoS website, hope the one in next box is acceptable. Please note that Ethical considerations in conducting a Cochrane review are available at: http://www.cochrane.org/editorial-and-publishing-policy-resource/ethical-considerations	Inserted at the end of Methods: "Ethics approval and patient consent forms are not provided as they are not necessary for a Cochrane review, of which this study is a product".
61.	6. Thank you for providing a link to the reports. We very much support the reports being made available and we ask that you deposit the information in DRYAD; http://datadryad.org/	Posting to Dryad requires We support Dryad too but cannot post the clinical study reports as soon as we have a doi.	For peer review purposes we set up a Dropbox access which is still extant: https://www.dropbox.com/s/8ofs5qyjmkd7bxk/Jefferson%20et%20al%20ROB%20paper%20manuscript%20-%20supporting%20files.zip Please note that our intention was stated in our original submission letter ("As a supplementary item we are making the source table of all risk of bias assessments available to you and your reviewers and are happy to share the relevant complete CSRs. This however will have to be through a drop box type system, given their size") and the url was passed on to your editorial

			office on the First of November.
62.	Academic Editor comments: The lack of clarity around which parts of the CSRs were analyzed at which stages of the study needs to be fixed. I'm also concerned about eliminating the "Unclear" category when evaluating complete study reports. The authors should provide a better rationale for that choice, and discuss the implications of that choice for the conclusions they can draw. Alternatively, they can include the "Unclear" category in their evaluations of CSRs and present that data in the revised paper;	Two issues are mentioned in this comment. The 3-stages issue: the convergence of editors and reviewers have led us to reanalyze our publications - core report – full report sequence. We realized that comparing core or full reports with publications across reviews is potentially confusing and misleading. Many of the trials included in the latter Cochrane reviews had no correspondent articles because of sizeable publication bias. The Cochrane Risk of Bias tool changed in 2010 making direct comparison of assessments difficult. Finally, publications reported far fewer items that could be assessed, leaving many of the boxes empty. The rationale for eliminating the "unclear" risk of bias category issue: this seems to us difficult to argue with, if one is to simply skim the clinical study reports we have made available. The reports form the scientific basis for approval of a pharmaceutical. Everything that should be of interest to regulators to assess whether the drug is better than placebo (including thousands of pages of individual listings) should be in the reports. Risk implies uncertainty, but reports are supposed to	We have simplified the stages throughout the manuscript by downsizing the first stage (publications) giving the reasons listed in the column to the left. For example in Results the following has been added: We could not carry out a comparison of risk of bias judgments of journal articles with core reports or full clinical study reports because our assessments were largely based on secondary and not primary publications of the trials and an outdated risk of bias tool. There were therefore too few studies for which we had distinct risk of bias judgments of primary journal publications (many studies for which we have clinically study reports were and remain unpublished). In addition, the current Cochrane risk of bias tool was introduced after the production of our review based on published articles, making the comparison, had we had the data to undertake it, more difficult to interpret and possibly unfair.

		provide certainty of full design and reporting of the trial. If an item is not reported bias is present. We realize that switching gears from publications to clinical study reports is difficult, but it is essential now as we move to a different evidence paradigm. As an example, if an item such as randomization is not reported in full, bias is present.	
63.	Reviewer Notes: Reviewer #1: I confine my remarks to statistical aspects of this paper. These were very simple; I think some tests could be usefully added and I have some other comments as well;	See serial 12	Nil
64.	General comment: Why is risk of bias categorized into "low", "unclear" and "high"? Since the tool used to assess the risk has 7 items, couldn't more precise levels of risk be determined;	Apologies, there is probably a misunderstanding. The Cochrane tool has 6 domains, with possibly more than one source of bias each of which is rated low/unclear/high. So the levels of risk are 3. It is impossible to list all possible sources of “other bias” and the tool does not list them.	The following text has been added to Methods: “The current Cochrane risk of bias tool was first introduced in 2010. The tool consists of six domains, each may have more than one source of bias application, depending on the subject matter.[7] Our applications were as follows: selection bias (random sequence generation and allocation concealment), performance bias (blinding of participants and personnel – all outcomes), detection bias (blinding of outcome assessment - all outcomes), attrition bias (influenza symptoms, complications and harms outcome data), reporting bias (selective reporting) and other bias. The identification of sources of other bias are left at the reviewers’ discretion.”

65.	line 109 ff Say more about how the Cochrane method works. The 7 items are in an appendix, they should be here, as should be a description of how they are combined;	See serial 9	See Serial 9
66.	line 140 ff Why does complete information mean that no trial will have an "unclear" amount of risk of bias;	See Serial 7	Nil
67.	line 151-2 I am certainly not a big advocate of p values, but statistical tests do have some purpose. They could be used here to distinguish the amount of risk of bias at different levels of information. A simple chi-square test would probably suffice. (And would doubtless be hugely sig.);	Serial 8 also refers	A simple chi-square test is not appropriate because we are (re) assessing the same trials based on differing amounts of information available In addition the purpose of conducting a hypothesis testing is to provide inferences to the wider population. However we are unsure if this is appropriate given our lack of experience with clinical study reports.
68.	line 165 Since the authors said that no complete report could be unclear, it is not sensible to report on the proportion that were unclear. It is 0 by definition;		Text has been clarified to: “compared to none using complete clinical study reports”
69.	Tables 1 - 3 would probably be clearer if the columns were ordered "high", "unclear", "low";		Tables 1 and 3 were removed. Table 2 (now 1) has been edited accordingly
70.	Figure 1 would probably be clearer as a mosaic plot (or two);	After removal of the publications ROB analysis from the review (see Serial 7) we played with several figure formats. None were satisfactory and the figure was removed	Figure removed

71.	Reviewer #2: Reviewer: Beate Wieseler	-	
72.	General comment: This study provides important and new information on the relevance of clinical study reports (CSRs). As such it is specifically important given the current initiatives to make CSRs publicly available. It furthermore addresses possible implications of availability of CSRs for systematic reviewers, which might start a timely discussion. However, some key issues would need to clarified before publication (e.g. unclear distinction between a "core" and a "full" clinical study report in the methods section as well as inconsistencies in sample sizes in the results section);	The difficulty is double for those not used to clinical study reports: Roche-speak ("Module 1) and ICH-speak (e.g. "core report").	In the introduction the following has been added …."For the purposes of this paper the core report plus all it appendices (roughly equivalent to modules II to V in oseltamivir clinical study reports) will be known as the full clinical study report"…. We also have gone through the text to ensure consistency of terms
73.	Abstract Lines 34-35: 11 articles, 15 CSRs: how many studies;		The text now reads: "We used and extended the Cochrane risk of bias tool to assess and compare risk of bias of 14 oseltamivir trials (reported in 10 clinical study reports) obtained from the European Medicines Agency (EMA) and its manufacturer, Roche."
74.	Lines 36 - 44: The results are difficult to understand, please also see comments on results section of manuscript		We have deleted references to 2009, 2012, 2013 etc and re-written the results. We hope this is now clear. We now refer to "journal publications", "core reports" and "full clinical study reports" in a consistent manner
75.	Line 43: It is unclear from the abstract how the results presented support the conclusions. From my point of view this only becomes clear		The conclusions have been edited

	from the discussion;		
76.	Introduction Line 69: I would like to suggest using "according to sections 1-15 …" because ICH E3 is not meant as a template of a CSR but describes required content (and the general structure);	ICH E3 is a guideline, and it would be wrong to suggest the wording 'according to ICH E3' in the text, as it is not mandatory for industry to follow	Nil
77.	Line 70-76: The first part of the section introduces a definition of a core CSR (report according to sections 1-15 of ICH E3); however, in the second part of the sentence appendices are included (which would not be part of a core report). This could easily be solved by splitting the sentence in two and clearly referring to appendices as part of a full CSR. The definition of a core and full CSR is also unclear from other parts of the manuscript (please see below);		The following had been added to the Introduction: "For the purposes of this paper the core report plus all it appendices (roughly equivalent to modules II to V in oseltamivir clinical study reports) will be known as the full clinical study report"
78.	Lines 78-85: It remains unclear from the text why funding timelines resulted in only 20 of 32 trials being included and why only core reports were included. This only becomes clearer in the methods section. As such I find this paragraph in the introduction rather difficult to understand. Please rephrase;		The para now reads: "In 2012, we published an update of our Cochrane review of neuraminidase inhibitors for which a total of 32 oseltamivir trials were eligible. Unlike most Cochrane reviews, this review was based only on clinical study reports but because of the lateness of delivery of clinical study reports and our funding timelines, the review update was based only on core reports."
79.	Lines 93-98: What is the research question; 1) Investigating the influence of progressively greater amounts of information in CSRs on the	The 3 objectives are summarised in the abstract phrase: "Here we analyze whether progressively greater amounts of information	The end of the Intro now reads: "In this report we describe our use of these tools

	risk of bias assessment or;	and detail in clinical study reports….”	to address three specific questions: 4. Do core reports change the risk of bias evaluation compared to published papers? 5. Do full clinical study reports change the risk of bias evaluation compared to core reports? 6. Do full clinical study reports change the risk of bias evaluation compared to published papers? In summary we intended to analyze whether progressively greater amounts of information and detail in clinical study reports (including trial protocols, statistical analysis plans, certificates of analyses, individual participant data listings and randomization lists) affected our risk of bias assessments”. We then explain in the results that objective 3 could not be achieved (see also serial 7)
80.	2) Investigating whether the Cochrane risk of bias tool plus the additional instrument can be used to assess risk of bias of trials reported in CSRs;		
81.	Methods General comments on methods: The study analyses the outcome of a risk of bias assessment based on journal publications, core CSRs and complete CSRs. Lines 68 to 70 of the manuscript define the core CSR as those parts of the CSR prepared according to sections 1-15 of ICH E3, i.e. the main report from the title page to the reference list but without the appendices. Using the Roche CSR structure this would be Module I only. From the methods section it remains unclear whether this definition of a core report was	On reflection we agree with comments.	We have taken the unsatisfactory comparison with risk of bias of journal publications and simplified the text comparing core reports with full reports for the 14 trials in 10 CSRs mentioned at the beginning of Methods. See also Serial 7

	also used for the analysis or whether any appendices (either received from EMA or provided by Roche, Module II according to Roche CSR structure) were also included in the assessment of risk of bias labeled as a core report in Tables 1 and 2. Please clarify. Please also provide a clear definition of core and full CSR in the methods section. Although there is no established definition in the literature, from my point of view the core report would be the part of the CSR according to sections 1-15 of ICH E3, i.e. without any appendices;		
82.	Please also consider the following two issues: 1) From the supporting material it seems that appendices (Module 2) were included in the "core report" in your analysis For example, in the Excel table provided as supporting information (CIST M2 table) line 36 for study WP16263 (element: blinding of participants) describes as the rationale for the 2012 assessment (which I understand is the assessment you used as "based on core CSR", lines 107-108 of the manuscript): "Placebo and oseltamivir capsules were described as having non-identical appearances from the certificate of analysis: oseltamivir: "Body: grey, opaque; cap: light yellow, opaque" placebo: "Body: grey, opaque; cap: ivory, opaque""; According to the definition of a core report referring to sections 1-15 of ICH E3, the certificates of analysis would not be part of the	We agree with the comments and have taken the action listed in serials 7 and 26. Placebo description was available in the text of one core report	

	core report because they are part of the appendices;		
83.	2) According to the Jefferson 2012 Cochrane Review (Table 9), more than Module 1 seems to have been available for at least 5 CSRs. Does that mean that more than Module 1 was used in the risk of bias assessment in the Jefferson 2012 review (and is the risk of bias assessment of Jefferson 2012 indeed presented as an assessment "based on core CSRs" in the manuscript?);	Jefferson 2012 used mostly Ms1, but also some Ms2 as they came in before timelock. For this paper, we're restricting our analysis to trials for which we only had M1 in Jefferson 2012. The order in which we received clinical study reports was outside our control. It could have introduced some bias, although we have no evidence of that.	This reflection has been added as a potential limitation of our study
84.	Lines 100-103: It is unusual that several studies are reported together in one CSR. Please explain in more detail (what were the reasons for this, was this justified, did the reports still include a full account of information on the individual studies?)		The following has been added to Methods: "The reporting of more than one trial in the same clinical study report is unusual. Roche gave low influenza circulation and the consequent need to pool studies as the reason."
85.	Line 104: Since there are a number of Cochrane reviews on neuraminidase inhibitors, please include the relevant citation: … timelock for our 2012 Cochrane review update [reference];		References have been added
86.	Lines 109-120: It does not become clear from this paragraph for how many studies core reports or core reports with (exactly which) appendices were available. Please clarify. An appendix table describing the available parts for each of the 15 reports might be helpful;	NB: **Reference 10** in the manuscript is a bookmark for the reference of our latest Cochrane review update which is currently undergoing peer review. We will insert the full reference as soon as it format is known.	We have revised the text to as follows: "In April 2011, we began to obtain the appendices of the clinical study reports included in our review. For most clinical study reports we

			requested, EMA had the protocol, protocol amendments, statistical analysis plan, blank case report forms, and other appendices contained in what Roche terms the second "module" of a full clinical study report (see Appendix 1). However EMA did not possess—and therefore could not provide us with—full clinical study reports with the exception of trial WP16263.[8] For approximately three years Roche had repeatedly refused our requests for full clinical study reports.[9] In the course of carrying out these new extractions, Roche changed its policy on access to data and pledged in April 2013 to share with us 74full clinical study reports (www.bmj.com/tamiflu/roche). Twenty trials were included in the analysis of our current Cochrane review.[10]. As we were already in possession of core reports and appendices such as the protocol and statistical analysis plan for the 14 trials in this analysis, the additional data for other clinical study reports provided by Roche does not concern this paper. In the Clinical study reports Roche redacted information that they judged to be of "legitimate commercial interest" or present a risk of trial participant re-identification. For our purposes, the redactions did not impede an analysis of risk of bias. Based on our growing familiarity with clinical study reports, we designed and piloted an extraction sheet to record how our understanding of the trials changed in light of

			availability of the additional appendices. “
87.	Lines 128-133: In contrast to lines 100 to 103 here you are referring to 74 CSRs. Please clarify the different numbers. Which CSRs (studies) were used in your study (only the 15 reports [covering 20 studies] mentioned in lines 100-103?);		Text has been clarified throughout the manuscript
88.	Lines 135-139: Which studies were included in your investigation presented in the paper;		See Methods and response to Serial 31
89.	Lines 140-146: For assessments of risk of bias based on full CSRs you did not allow an "unclear" judgment. This seems to be justified. However, since e.g. unclear allocation concealment is not necessarily inappropriate allocation concealment, you might be underestimating the quality of the trials. This has implications for the interpretation of your results, which does not become fully clear from your results text and discussion. Please address this issue (please also see comments below);	We think the reviewer is taking the risk of bias judgments as 100% objective. All such judgments are challengeable. If allocation was unclear, we would have said “high” risk of bias, because of the logic of our judgment already explained. When the originals trials were designed, allocation concealment would not have been seen as a source bias. Initial publications that revealed its importance only came to light in 1995 and its incorporation into CONSORT was much later than the trial ending (JAMA. 1995 Nov 8;274(18):1456-8.Subverting randomization in controlled trials. Schulz KF et al)	Nil
90.	Lines 149-150: It does not become clear how the extraction of risk of bias assessments described in lines 149 to 150 (citing the 2010 [Jefferson et al] and the 2012 [Wang et al] Cochrane review) relates to lines 107-108 (citing the 2012 Jefferson et al review). Please clarify. It might be meaningful to describe all	Extraction is too technical a term	Edited to: “we used” instead of “we extracted”

	data extractions together in one paragraph;		
91.	Results General comment on results: The timing of risk of bias assessments in the various versions of the Cochrane reviews and the documents on which these assessments were based does not become clear from the manuscript without going back to the cited Cochrane reviews and the supplementary materials. In addition, the number of studies and publications, core reports and full CSRs used is unclear from the text and tables. Specifically, it is not clear why 11 core reports are used in Table 1, 15 in Table 2 and 11 in Table 3. Please clarify. It might be helpful to provide a flowchart or some other sort of graphical representation;		See responses at serials 7 and 26
92.	General comment on tables: 11 core reports are used in Table 1, 15 in Table 2 and 11 in Table 3. This probably is due to the fact that you are trying to include the maximum of available risk of bias assessments in your analysis. As an additional analysis, it might be useful to perform all comparisons on a consistent sample of trials/documents. This would allow describing information gain along the line of adding more and more parts of CSR documents in a given sample of trials (please see also comment on lines 159-161 below);	Good suggestion	We have restricted our comparisons to 14 trials throughout (core vs full study reports) See also responses at serials 7 and 26
93.	Tables 1 to 3: Please explain how the total number of judgments is derived in the tables	We agree	Only 1 Table in the manuscript now

	(89 in Table 1, 130 in Table 2, 90 in Table 3). I assume that, in addition to the different number of core reports included, the differences are due to different numbers of outcomes assessed. However, it remains unclear why you have a different number of judgments in the publications sample in Table 1 (89) and Table 3 (90);		
94.	Table 2: Compared to the judgments from core reports, adding information from complete CSRs did not change any "high" judgments. According to your methodology you do not accept any "unclear" judgment any more at this stage. Therefore, any "unclear" judgments that cannot be solved from complete CSRs change into "high" judgments. It would be interesting to know which part of the "high" judgments is due to additional information from complete CSRs leading to informed "high" judgments and which part is due to still unclear information;	Good point. NB: **Reference 10** in the manuscript is a bookmark for the reference of our latest Cochrane review update which is currently undergoing peer review. We will insert the full reference as soon as it format is known.	We have inserted a sensitivity analysis in the text of results: “Had we kept the unclear risk of bias judgment in our current review [10] we would have had 64 unclear occurences. The breakdown of these 64 into the various attributes is: Attrition bias: symptoms (10); complications (9); safety (15) – these are unclear because we do not know the impact of missing symptoms data; unclear definitions for complications; and compliharms Other bias (13) – these are unclear due to the unknown effect of the de-hydrochloric acid Performance bias (6) – these are unclear due to the missing certificate of analysis describing the placebo appearance Selection bias (10) – these are unclear due to the missing or unclear randomisation lists meaning we cannot confirm random sequence generation

			Detection bias (1) – unclear due to unknown impact of different coloured placebo caps on outcome assessment” We have inserted two new tables reporting the results of a sensitivity analysis allowing unclear judgments in both core and full clinical study reports (Table 2 and 3).
95.	Figure 1: Please clarify in the figure legend why the columns for 2012 and 2013 do not present all low, unclear and high judgments together;		Figure 1 has been deleted
96.	Lines 158: Table 1 includes 74/89 "unclear" judgments for journal publications, not 75/90, please clarify;		74/89 is correct
97.	Lines 159 - 161: I don't think one can follow the "unclear" judgments from Table 1 to Table 2 because there are different samples in the two tables (you might want to add a set of tables based on the same sample to allow for such a follow-up from publication to core report to full report). It would be interesting to provide examples of changes from unclear to low to high describing the information which led to these changes in judgment;	.	See serial 39, but remember the judgments are subjective
98.	Lines 162-167: The fact that there are 0 judgments of unclear risk of bias based on full		See serials 39 and 42 and clarifications inserted

	CSRs is (just) a consequence of your definition. Please clarify this in the text of the results section; otherwise the reader might understand that full CSRs provided full information in all cases. Furthermore, please clarify how many of the "high" judgments are due to still unclear information and how many are due to the assessment of the available information in the full CSRs;		in text
99.	Lines 154-167: It would be interesting to know in which of the elements of the risk of bias assessments the most prominent changes were seen.Is it possible to at least qualitatively describe what type of information led to the changes;		See serials 39 and 42 and clarifications inserted in text
100	Discussion Lines 184-185: ""Unclear" risk of bias often became a more certain "low" or "high" risk of bias, or even certainty of bias or certainty of absence of bias. I understand that more complete information from a full CSR can result in certainty of bias or certainty of absence of bias. However, from my point of view it is open for discussion whether there is a higher risk of bias from the fact that there is missing information from a full CSR versus a core CSR or a (full or core) CSR versus a journal publication. E.g. with examples 1and 2 from Box 1: missing information from the core CSR leads to unclear risk of bias while missing information from a		See serials 39 and 42 and clarifications inserted in text

	full CSR leads to high risk of bias. In your methods section you define missing information from a CSR as high risk of bias. You might consider discussing this in more detail;		
101	Lines 186-187: "Certainty or low levels of uncertainty are due to our expectations regarding the complete clinical study reports." I understand you are expecting certainty or low levels of uncertainty from full CSRs. However, I do not understand what this sentence means.;	Thank you	Text changed to: ” When the information was not available, our judgments changed because we found gaps in the availability of information and inconsistent information. Whether the full study reports represent an exhaustive and coherent source of trial narrative and data remains unclear”.
102	Potential additional points for the discussion: You correctly point out that with full CSRs there is the possibility of understanding the timeline of a study concerning study planning, enrolment and treatment of patients and analysis and reporting of data. According to my experience, there might be changes to the protocol after the start (or even the end) of enrolment (but before unblinding). I am not sure this always results in a high risk of bias (in blinded studies). As with the availability of full CSRs, systematic reviewers would use this information in their risk of bias assessment, there is a need to discuss under which circumstances this should result in a high risk of bias and in which circumstances this would be less critical. The same is true for the development of a full statistical analysis	The reviewer appears to regard risk of bias assessments as objective and assigned on the basis of pre-set scenarios. This is not so especially since this is the first time to our knowledge that such assessments are made on clinical study reports. We find it very difficult to speculate on specific scenarios of levels of bias. This is why we agreed to regard any unclear item in full study reports as unclear.	Nil

	plan (for blinded studies). While obviously the crucial issues and general lines of analysis should be pre-defined in the protocol, often detailed analysis plans are prepared while the study is ongoing (but before unblinding). This need for clarification (and other open questions resulting from access to full CSRs for risk of bias assessment) could be an additional point for the discussion.;		
103	I am wondering if the discussion could also address earlier work on the relevance of study protocols for risk of bias assessment and what full CSRs would add. For example a paper by Soares et al. addresses information gain from protocols on study methods: Soares H et al. Bad Reporting does not mean bad methods for randomized trials: observational study of randomized controlled trials performed by the Radiation Therapy Oncology Group. BMJ 2004; 328:22 A commentary on the Soares paper is also addressing the impact for systematic reviews: Del Giblio A et al. Commentary: The quality of randomized controlled trials may be better than assumed. BMJ 2004;328:24 We have quantified information gain from CSRs for some methods items used for risk of bias assessments: Wieseler et al. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports and journal publications. BMJ 2011;	We are not sure the Soares paper is an appropriate example as it compared methodological features in protocols and published counterparts in a population of 56/59 published trials. We have already referred to Wieseler's later study (ref 20): Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervölgyi V, Kohlepp P, et al. 395 Completeness of Reporting of Patient-Relevant Clinical Trial Outcomes: Comparison 396 of Unpublished Clinical Study Reports with Publicly Available Data. PLoS Med. 2013 397 Oct 8;10(10):e1001526.;	Nil

	344:d8141;		
104	Comment on supporting materials Excel table provided as supporting information (CIST M2 table): The study IDs given in the table do not match the study numbers given in the methods section (which probably are report numbers). Please clarify by providing both study IDs and report numbers in the appendix table.;	The Table reports risk of bias assessments for all trials in the reviews, not just the ones in this study	Nil
105	Reviewer #3: This is an excellent summary of the changes in the risk of bias comparing the data in published articles in journals to clinical study reports. The information is timely and useful, and points to the need for examining all the evidence when evaluating medical interventions.;	Thank you	
106	A few comments that would aid in clarity of the manuscript: * The number of actual clinical trials vs the number of complete study reports is not clearly explained up front and it take a while to gather that there are multiple studies included in the various study reports. Would be helpful to clearly explain this in the abstract and introduction;	We agree	See serials 7 and 26
107	* The point about the Cochrane risk of bias tool not being optimal is an important one. The Cochrane tool has a broad category called "other risk of bias". Such a broad category is often not helpful in examining data as it relies	We agree	See serial 9

	on the reviewers personal knowledge of types of bias or actually looking for various types of bias. The seminal paper by Sackett listed over 30 types of bias that could be present in various types of studies (both randomized and observational) and represents a more comprehensive way of evaluating bias;		
108	* One major issue is the authors eliminating the category of "unclear" risk of bias when evaluating complete study reports. While it is agreed that there should be no information left out of these reports, the unfortunate truth is that even with these more thorough sources of information, sufficient detail is often missing. For instance in Box 1 the authors point out that insufficient detail is available regarding issues like the randomization code. Therefore there is still "unclear" risk of bias when information is missing. Eliminating the "unclear" category has two consequences: first it means that the published and the complete study reports are not judged by the same standard which could itself bias the assessments of bias; second, it assumes that absence of evidence if evidence of absence - because detail is not present then it must not have been done or been done wrong. This is in unverifiable assumption. It would be better to reanalyze the results and use the "unclear" category when insufficient detail is present in CSRs. This is still a major problem in terms of trial transparency since as the authors state there is no reason these	We agree	See responses to serials 7 and 39

	important details should be withheld.;		
109	* The idea that missing data results in actual bias rather than "risk of" bias is an interesting notion. However in some situations when analyzing different imputation methods the best and worse case scenarios give the same qualitative (if not quantitative) results so that the bias does not affect the robustness of the conclusions. However the authors point is well taken that in many cases real bias exists and it more than just "potential" bias.;	Thank you	Nil
110	* It would be helpful if the authors could make some suggestions as to when bias is so limiting as to make the conclusions of meta-analyses unreliable. It is unfortunate to see meta-analyses in the literature of studies with high risk of bias, but when the results of these meta-analyses are presented the risk of bias is not considered or ignored, and the conclusions are presented as confirmatory.	We agree with the reviewer. We think when bias is so limiting as to make a meta-analysis unreliable, it either should not be done or should be done alongside a prominent explanation of the extreme limitations.	The following text has been inserted in Discussion: “We suggest that when bias is so limiting as to make meta-analysis results unreliable, it either should not be done or a prominent explanation of its clear limitations should be posted alongside the meta-analysis.”